Rapid C4I High Performance Computing for Hyperspectral Imaging Exploitation

George O. Ramseyer, Scott E. Spetka*, Richard E. Linderman and Brian C. Romano

Air Force Research Laboratory/Information Directorate
26 Electronics Parkway
Rome, NY 13441-4514
Phone: (315)330-3492
George.Ramseyer@rl.af.mil

*SUNY Institute of Technology
PO Box 3050
Utica, NY 13504
scott@cs.sunyit.edu

Abstract

A multidisciplinary effort has been initiated which spans the C4I and signal/image processing computational technology areas to integrate diverse capabilities of existing hyperspectral image exploitation systems. The primary objective of the project is to develop and demonstrate support for the rapid, low latency exploitation of hyperspectral information. A flexible and open framework will service exploitation requests from C4I users by tapping into large, dynamic databases of previously processed and raw data to deliver products to the requester with minimal latency. A web-based interface is being developed so that any authorized user with a web browser can input requests. The user can select data sources, exploitation time intervals, and a parallelized exploitation method for execution. To maximize productivity and minimize the decision making cycle of the requestors, algorithm parallelization efforts seek to achieve a minimal latency before initial results begin to stream back to the requestor in typical web prioritized fashion.

1. Introduction

The research presented here is a key ingredient in facilitating the rapid exploitation of hyperspectral imaging data as part of a multi-dimensional application of *information with force* to seize the initiative in the battlespace. To rapidly seize the initiative, it is necessary to effectively integrate Intelligence, Surveillance and Recognizance (ISR) with operational Command and Control (C2) into C4I [Ackerman], with computing power being a key enabler. *Global Awareness* is the Air Force's answer. The goal of Global Awareness is to provide ubiquitous, consistent and integrated battlespace information on demand that is tailored to the needs of the Commander and the Warfighter [Phister].

Report Documentation Page

Form Approved
OMB No. 0704-0188

Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.

1. REPORT DATE **JUN 2001**	2. REPORT TYPE	3. DATES COVERED **00-00-2001 to 00-00-2001**

4. TITLE AND SUBTITLE **Rapid C4I High Performance Computing for Hyperspectral Imaging Exploitation**	5a. CONTRACT NUMBER
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Air Force Research Laboratory,Information Directorate,26 Electronics Parkway,Rome,NY,13441-4514**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES
6th International Command and Control Research and Technology Symposium

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**		**7**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

The first section of this paper describes how differences in requirements distinguish the DOD approach from other efforts by NASA and the private sector. Then, we describe existing system components that are used in developing the architecture of an integrated hyperspectral image exploitation system. The next section addresses developing technologies and discusses their role in motivating and enabling the system development that we have undertaken. The next section introduces the architecture of the central system component that is used to integrate existing technologies and is supported by the latest developing technologies. We present the architecture and introduce issues that may impact the design and development of the system. The paper concludes with a discussion of future work.

2. DOD Requirements

The two-decade investment by DoD, NASA and private industry into the development of hyperspectral imaging has resulted in the transformation of hyperspectral imaging from a bench-top scientific curiosity to a discipline with fielded demonstration systems. A single hyperspectral image consists of two hundred or more bands, times the number of spatial pixels in the horizontal and vertical direction, yielding data acquisition rates of tens to hundreds of megabytes per second. Most of the data that will be collected will be stored in ground-based data repositories. High performance computing can play a key role by supporting timely exploitation of these voluminous data sources.

Although DOD, NASA and commercial approaches to hyperspectral image exploitation have a great deal in common; specific DOD requirements are also recognized. DOD requirements for hyperspectral image exploitation are distinguished by the critical need to rapidly process new as well as previously acquired raw hyperspectral data, so that a diverse and distributed community of intelligence analysts and battlefield decision makers can take appropriate actions, based upon these analyses, in near real-time.

3. Existing System Components

3.1 Hyperspectral Data

State-of-the-art hyperspectral sensors generate large amounts of data at rates on the order of 10's of Mbytes/second. This results in large files on the order of 10's of Gbytes for individual scenes covering suitable areas [Cooley]. These data sets consist of many traditional images, each acquired over a relatively narrow electromagnetic bandwidth and co-registered – forming a 3-dimensional "image cube". A single hyperspectral sensor collects 100's of contiguous bands in the Visible/Near-IR, Short-Wave IR, Mid-Wave IR or Long-Wave IR regions of the spectrum.

The processing of these image cubes is highly redundant, requiring the same processing be performed on hundreds of thousands of spectrums in a "spectral filter" configuration for a single image cube. This kind of image cube processing lends itself readily to being data parallelized both across the spatial and spectral dimensions of the image. However, it currently takes a considerable amount of time to serially process hyperspectral data.

The parallelization components of this project will greatly reduce the time delays of processing of the ISR information so that it can be sent to C2 decision-makers.

3.2 Parallelization of Hyperspectral Algorithms

Scalable and portable algorithms are being parallelized at the Air Force Research Laboratory's Space Vehicle (VS) and Sensor (SN) Directorates. The Space Vehicle Directorate is parallelizing atmospheric correction algorithms, which correct raw hyperspectral data for atmospheric conditions. The Sensor Directorate is parallelizing code that exploits *a priori* knowledge of objects.

The Space and Warfare's (SPAWAR's) System Center is parallelizing algorithms that were developed by the Defense Advanced Research Projects Agency (DARPA) sponsored Adaptive Spectral Reconnaissance Program (ASRP). Also, the Naval Research Laboratory's (NRL's) Optical Real-time Analysis and Spectral Identification System (ORASIS) algorithm suite, which relies on pattern recognition instead of band combinations or statistical techniques, is being parallelized. The Army Research Laboratory (ARL) is parallelizing code that is being used for target detection and clutter rejection.

3.3 Broadsword

Broadsword is a Web-based system that provides C2 and ISR with the ability to meet the goal of striking a target within seven minutes [Lawlor]. The key to this is faster and more accurate responses in the battlespace to information that exists in all forms and resides in databases throughout the military. Broadsword allows users to search, find and retrieve data from a variety of heterogeneous and diverse sources with an associated security infrastructure. Broadsword consists of five functional components: the gatekeeper, the keymaster, the access and authentication module, the trusted transfer agent and the clients.

4. Developing Technologies

4.1 High Performance Computing

Developing software for scalable HPC systems remains technically challenging. The Common High Performance Computing Software Support Initiative (CHSSI) develops efficient, scalable, portable software codes, algorithms, tools, models and simulations that run on a variety of HPC platforms and are needed by a large number of science and technology (S&T) and test and evaluation (T&E) scientists and engineers. CHSSI helps the DoD take advantage of existing and future computing and communications capabilities by building software with an emphasis on reusability, scalability, portability and maintainability. (http://www.hpcm.dren.net/Htdocs/CHSSI/index.html).

There are four Major Shared Resource Center's (MSRC's) that currently operate large high performance computer systems that are available to the entire DoD HPC

community. Each MSRC has been designated to provide a complete HPC environment, data storage, and include a full range of resources including hardware, software, data storage, archiving, visualization, training, and expertise in specific computational technology areas. http://www.hpcmo.hpc.mil/Htdocs/MSRC/index.html Also, Distributed Center's (DC's) have been established to provide HPC capability to a specified local and remote portion of the HPC community. These are typically modest-sized systems where there is a significant advantage to having a local HPC system where there is a potential for advancing DoD applications using investments in HPC capabilities and resources. http://www.hpcmo.hpc.mil/Htdocs/DC/index.html

4.2 Cluster Computing

The increased importance of integrating access is being driven by the increasing numbers of inexpensive computer clusters that are currently available. Advances in technology in three areas make the vision of integrated hyperspectral image exploitation possible. Improvements in application portability, based on compiler technologies and opensource freeware library packages, such as MPI [http://www-unix.mcs.anl.gov/mpi/mpich/], facilitate implementations on heterogeneous platforms. Improvements in operating system support help to harness clusters of inexpensive processors, such as the FreeBSD-based BUDS (BSD Unix Distributed Simple-ly) project (http://www.mozie.org/projects/buds/index.html). The Linux systems support clusters [Beowulfhttp://sdcd.gsfc.nasa.gov/ESS/annual.reports/ess97/sys/linux.html] as does the Compaq Tru64 Unix cluster system, formerly known as Digital Unix clustering [http://tru64unix.compaq.com/faqs/publications/pub_page/cluster_list.html]. Projects like NGI [www.ngi.org] and Internet2 [www.internet2.org] provide improved access to supercomputing centers at network speeds of 600Mbps.

4.3 Joint Battlespace Infosphere

The Joint Battlespace Infosphere (JBI) is an emerging combat information management system that provides individual users with the specific information required for their functional responsibilities during crisis or conflict. The JBI integrates data from a wide variety of ISR sources, aggregates this information, and distributes the information in the appropriate form and level of detail to C2 users at all echelons. The JBI, through intelligent agents, queries existing databases and creates databases on demand, and also provides the opportunity to exchange information between various C2 systems and databases. [USAF Scientific Advisory Board].

5. The Hyperspectral Image Exploitation "FrameWork"

A "FrameWork" for driving a hyperspectral image exploitation system is being implemented to integrate diverse DOD high performance computing systems that are geographically distributed. These DOD algorithms have been developed at a significant cost but they are currently, in many cases, difficult to access so analysts are often limited

in the algorithms at their disposal. The FrameWork will potentially present a unified interface to all DOD hyperspectral data and processing algorithms.

The FrameWork, shown in the diagram below, will use web technology to provide an interface that is globally available and easily maintainable. It will be installed along with a web server to allow access through a conventional browser. This avoids the cost of initial installation and upgrades for client workstations. An object-oriented implementation for the FrameWork will provide the ability to introduce new hyperspectral image types into the system and to introduce new algorithms for hyperspectral image processing. Raw data and processed hyperspectral image products can be supplied as input for the algorithms as well as any other inputs that are defined according to the algorithms introduced into the system. Processed images can be "published" through either the JBI interface or through Broadsword for DOD-wide distribution.

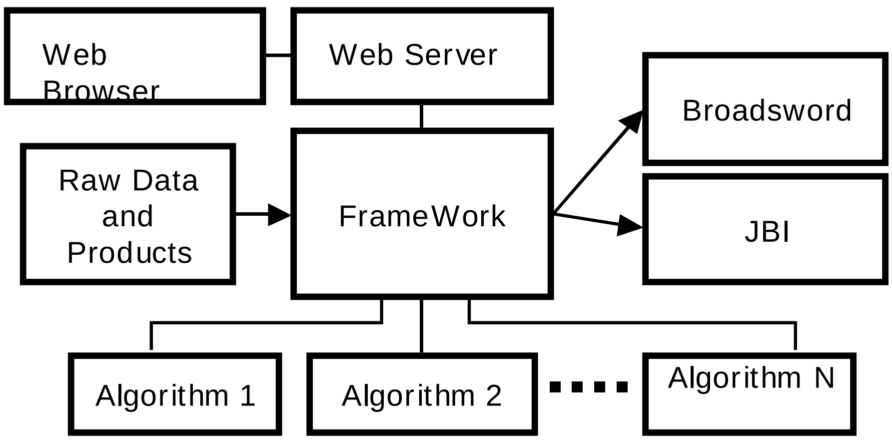

Fig. 1 - FrameWork Architecture

There are several issues to resolve in the framework implementation. Initially, the architecture will support algorithms that can be ported to a cluster-based system and executed directly from the FrameWork. Subsequently, the system will be extended to support remote execution of algorithms on remote platforms, such as the Intel Paragon. Many algorithms are operated through graphical interfaces that provide processing controls including image intensity, frequency ranges, and many other algorithm dependent controls. Appropriate values can be stored as *preferences* for users accessing the framework to facilitate access to remote heterogeneous data types and algorithms. The interface between the FrameWork and each of the algorithms must be capable of supplying all parameters necessary to operate the algorithm.

More complex processing models are also anticipated. Once hyperspectral data is widely accessible through high-speed networks, we expect a demand to emerge for both real-time applications and very complex batch mode computations, involving several data sources. A number of issues similar to those associated with distributed database query processing must be addressed. Other issues, including synchronization, fault-tolerance,

and redundancy, must also be addressed. Data movement optimization, to minimize latency and optimize algorithm execution, depends on the specific algorithm, space/time considerations and the style of parallelism that it implements. For example, algorithms that achieve parallel execution through pipelining may be able to operate at the end of a data stream through the network.

The goal is to enable the rapid, low latency exploitation of hyperspectral imagery on high performance computers (HPC's) in a flexible and open framework. The HPCs will service exploitation requests from users around the *Infosphere* by tapping into large, dynamic databases of previously processed and raw imagery to deliver products to the requester with minimal latency. The system will allow users to select data sources and parallelized exploitation methods for execution. An exploitation time interval will also be specified to further constrain algorithm execution. To maximize productivity and minimize the decision making cycle of the requestors, algorithm parallelization efforts will seek to achieve a latency of less than one minute before initial results begin returning to the requestor in typical web prioritized fashion.

6. Future Work

Future work includes the development of a Java-based system that will allow the integration of exploited hyperspectral imaging data into the system through the JBI [Holzhauer]. The web-like interfaces to publish, subscribe, and query core services will be drawn to the C4I community. To support this concept, technology and architecture are now being developed which supports globally distributed, tailorable and consistent actionable information. Key enablers include publish-subscribe information exchange paradigm, fuselets transform data into knowledge, and force templates. The JBI will enable Warfighters, C2 and ISR to customize and tune their information management enterprise to match force structures.

7. References

Ackerman, R.A., "Intelligence Enters the Fray", *SIGNAL*, pp. 31-34 (Feb 2001).

Cooley, T., "Modifying a Commercial Satellite for Military Applications", *Technology Horizons*, pp. 14-15 (Sept 2000).

Holzhauer, D.J., Ramseyer, G., Duncomb, R., Young, D. and Das, D. "Chemical/Biological Plume Analysis Knowledge Source (CPAKS)", *Proceedings: Command and Control Research & Technology Symposium*, (2000).

Lawlor, M., "Software System Slices Search Time", *SIGNAL*, pp. 25-28 (Feb 2001).

Phister, W.P. and Plonisch, I., “C2 of Space: The Key to Full Spectrum Dominance”, *Proceedings: Command and Control Research & Technology Symposium*, *Volume 2*, pp. 1254-1271 (1999).

United States Air Force Scientific Advisory Board, “Building the Joint Battlespace Infosphere” Vol. 1 Summary, SAB-TR-99-02 Dec. 17, 2000.

8. Acknowledgement

This effort is being funded by the Department of Defense High Performance Computing Modernization Program’s Common High Performance Computing Software Support Initiative (CHSSI).